IDENTIFICATION AND CLASSIFICATION OF PLANT LEAF DISEASE

V. Sathish[1], Dr. K. Ramesh Kumar[2]

[1]M.Tech Scholar, [2]Associate Professor,

Department of Information Technology, Hindustan University (India)

ABSTRACT

This paper presents a method for identifying plant leaf disease based on color. Agrarians are suffering from the issue rising from different types of plant diseases. Sometimes biologists are also unable to identify the disease that leads to lack of identification of right type of disease. The goal of proposed work to diagnose the disease using image processing and artificial intelligence techniques on image of plant leaves disease. The input image of leaves is converted RGB to HIS. Then leaf disease segmentation is done using K-means clustering. After segmentation the mostly green color pixels are masked based on specific threshold values that are computed using Otsu's method. Neural network is trained for classification.

Keywords: HIS, Image processing, K-means, Leaf disease, Neural Networks.

I. INTRODUCTION

India is an agricultural country. Farmers have wide range of diversity to select suitable fruit and vegetable crop. Research work develops the advance computing system to identify the diseases using infected images of various leaf spots. Images are captured by digital camera mobile and processed using image growing, then the part of the leaf sport has been used for the classification purpose of the train and test. The technique evolved into the system is both Image processing techniques and advance computing techniques.

Image Analysis Can Be Applied For The Following Purposes:

1. To detect diseased leaf, stem, fruit.
2. To quantify affected area by disease.
3. To find the boundaries of the affected area.
4. To determine the color of the affected area.
5. To determine size & shape of leaf.
6. To identify the Object correctly.

Disease management is a challenging task. Mostly diseases are seen on the leaves or stems of the plant. Precise quantification of these visually observed diseases, pests, traits has not studied yet because of the complexity of visual patterns. Hence there has been increasing demand for more specific and sophisticated image pattern understanding.

Various Types of Leaf Spot Diseases:

- Bacterial
- Fungal
- Viral

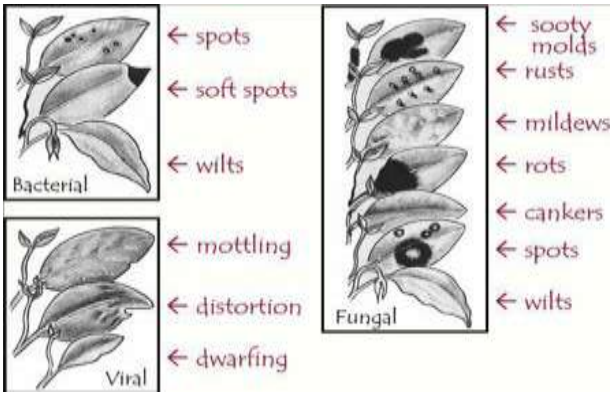

Figure 1: Various types of diseases

Most leaf diseases are caused by fungi, bacteria and viruses. Fungi are identified primarily from their morphology, with emphasis placed on their reproductive structures. Bacteria are considered more primitive than fungi and generally have simpler life cycles. With few exceptions, bacteria exist as single cells and increase in numbers by dividing into two cells during a process called binary fission viruses are extremely tiny particles consisting of protein and genetic material with no associated protein. In biological science, sometimes thousands of images are generated in a single experiment. There images can be required for further studies like classifying lesion, scoring quantitative traits, calculating area eaten by insects, etc. Almost all of these tasks are processed manually or with distinct software packages. It is not only tremendous amount of work but also suffers from two major issues: excessive processing time and subjectiveness rising from different individuals. Hence to conduct high throughput experiments, plant biologist need efficient computer software to automatically extract and analyze significant content. Here image processing plays important role [1].

In this project, we develop software for the automatic identification & classification of plant leaf diseases. Here the end-user is the farmer. This project classifies the plant leaves diseases hand into infected and non-infected classes. The developing software provides a fast and accurate method in which the leaf diseases are detected and classified using k-means based segmentation and neural networks based classification.

II. METHODOLOGY

The Methodology for diagnosing leaf diseases involves several tasks, such as Image acquisition, image enhancement, segmentation, feature extraction and leaf disease classification based on neural network.

Figure 2: Stages for classification of plant leaf diseases

Algorithm 1. Basic steps describing the proposed algorithm:

1. RGB image acquisition
2. Create the color transformation structure
3. Convert the color values in RGB to the space specified in the color transformation structure
4. Apply K-means clustering
5. Masking green-pixels
6. Remove the masked cells inside the boundaries of the infected clusters
7. Convert the infected (cluster / clusters) from RGB to HSI Translation
8. SGDM Matrix Generation for H and S. (Another name for gray-level co-occurrence matrix is gray-level spatial dependence matrix.)
9. Calling the GLCM function to calculate the features
10. Texture Statistics Computation
11. Configuring Neural Networks for Recognition

2.1 Image Acquisition

The different types of commercial crops, food grain, fruits and cereals samples both healthy and unaffected agriculture/horticulture produce used in the present work are collected.

2.2 Pre-Processing

Usually the images that are obtained during image acquisition may not be suitable straight for identification and classification purposes because of certain factors, such as noise, lighting variations, climatic conditions, poor resolutions of an images, unwanted background etc.

2.3 Color Transformation Structure

First, the RGB images of leaves are converted into Hue Saturation Intensity (HSI) color space representation. The purpose of the color space is to facilitate the specification of colors in some standard, generally accepted way. HSI (hue, saturation, intensity) color model is a popular color model because it is based on human perception. Hue is a color attribute that refers to the dominant color as perceived by an observer. Saturation refers to the relative purity or the amount of white light added to hue and intensity refers to the amplitude of the light. Color spaces can be converted from one space to another easily. After the transformation process, the H component is taken into account for further analysis. S and I are dropped since it does not give extra information.

2.4 Masking Green Pixels

In this step, we identify the mostly green colored pixels. After that, based on specified threshold value that is computed for these pixels, the mostly green pixels are masked as follows: if the green component of the pixel intensity is less than the pre-computed threshold value, the red, green and blue components of the this pixel is assigned to a value of zero. This is done in sense that the green colored pixels mostly represent the healthy areas of the leaf and they do not add any valuable weight to disease identification and furthermore this significantly reduces the processing time [2].

2.5 Removing The Masked Cells

The pixels with zeros red, green, blue components were completely removed. This is helpful as it gives more accurate disease classification and significantly reduces the processing time.

2.6 Segmentation

K-means clustering is used to partition the leaf into four clusters in which one or more clusters contain the disease in case when the leaf is infected by more than one disease. The K means clustering algorithms tries to classify objects (pixels in our case) based on a set of features into K number of classes. The classification is done by minimizing the sum of squares of distances between the objects and the corresponding cluster or class centroid [3].

2.7 Feature Extraction

The method followed for extracting the feature set is called the color co-occurrence method or CCM method. It is a method, in which both the color and texture of an image are taken into account, to arrive at unique features, which represent that image.

2.8 Color Co-Occurrence Method For Texture Analysis

The image analysis technique selecting for this study is the CCM method. The use of color image features in the visible light spectrum provides additional image characteristics features over the traditional grey scale representation.

The CCM methodology consists of three major mathematical processes. First, the RGB images of leaves are converted into HSI color space representation. Once this process is completed, each pixel map is used to generate a color co-occurrence matrix, resulting in three CCM matrices, one for each H, S and I pixel maps. Hue Saturation Intensity (HSI) space is also a popular color space because it is based on human color perception. Hue is generally related to the wavelength of a light and intensity shows the amplitude of the light. And saturation measures the colourfulness in HSI space.

Color spaces can easily be transformed from one to another. Following equations can be used to transform the images from RGB to HSI.

$$Intensity\ (I) = \frac{R+G+B}{3}$$

$$Saturation\ (S) = 1 - \frac{3\,min\,(R,G,B)}{(R+G+B)}$$

$$Hue\ (H) = 2 - ACOS\left\{\frac{[(R-G)+(R-B)]}{\sqrt[2]{(R-G)^2+(R-G)(G-B)}}\right\},\ B > G$$

$$Hue\ (H) = ACOS\left\{\frac{[(R-G)+(R-B)]}{\sqrt[2]{(R-G)^2+(R-G)(G-B)}}\right\},\ B \leq G$$

2.9 Neural Networks For Recognition

In this paper, neural networks are used in the automatic detection of leaves diseases. Neural network is chosen as a classification tool due to its well-known technique as a successful classifier for many real applications. The training and validation processes are among the important steps in developing an accurate process model using NNs. The dataset for training and validation processes consists of two parts; the training feature set which are used to train the NN model; while a testing features sets are used to verify the accuracy of the trained NN model. Before the data can be fed to the ANN model, the proper network design must be set up, including type of the network and method of training. This was followed by the optimal parameter selection phase. However, this phase was carried out simultaneously with the network training phase, in which the network was trained using the feed-forward back propagation network. In the training phase, connection weights were always updated until they reached the defined iteration number or acceptable error. Hence, the capability of ANN model to respond accurately was assured using the Mean Square Error (MSE) criterion to emphasis the model validity between the target and the network output [2].

III. CONCLUSION

This work consists of identifying the affected part of the plant leaf disease. Most affected part was segmented efficiently according to color classification tasks. Initially Image often more complex structures are needed in order to segmentation is done, and finally image analysis and make an optimal separation, i.e., correctly classify important features are extracted and classification performed using Neural network. An extension of this work will focus on developing hybrid algorithms such as genetic algorithms and NNs in order to increase the recognition rate of the final classification process underscoring the advantages of hybrid algorithms; also, we will dedicate our future works on automatically estimating the severity of the detected disease.

REFERENCES

[1]. Arti N. Rathod, Bhavesh Tanawal, Vatsal Shah, "Image processing Techniques for Detection of Leaf Disease" *International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 11, November-2013*

[2]. S. Ananthi and S. Vishnu Varthini., "Detection And Classification Of Plant Leaf Diseases" *IJREAS Volume 2, Issue 2 (February 2012) ISSN: 2249-3905*

[3]. Al-Bashish, D., M. Braik and S. Bani-Ahmad, 2011. "Detection and classification of leaf diseases using Kmeans- based segmentation and neural-networks-based classification" *Inform. Technol. J., 10: 267-275. DOI: 10.3923/itj.2011.267.275*

[4]. Ali, S. A., Sulaiman, N., Mustapha, A. and Mustapha, N., (2009). "K-means clustering to improve the accuracy of decision tree response classification" *Inform. Technol. J., 8: 1256-1262. DOI: 10.3923/itj.2009.1256.1262*

[5]. Applications of image processing in biology and agriculture J. K. Sainis, Molecular Biology and Agriculture Division,R. Rastogi, Computer Division, and V. K. Chadda, Electronics Systems Division, BARC newsletter.

[6]. Jayamala K. Patil, Raj Kumar, "Advances in image processing for detection of plant diseases" *Journal of Advanced Bioinformatics Applications and Research ISSN 0976-2604 Vol 2, Issue 2, June-2011, pp 135-141.*

[7]. Savita N. Ghaiwat, Parul Arora, "Detection and classification of plant leaf diseases usinf image processing techniques : A Review" *International journal of Recent Advance in Engineering & Technology(IJRAET) ISSN: 2347-2812, Volume-2, Issue-3, 2014*

[8]. Manoj Mukherjee, Titan pal, Debabrata Samanta "Damaged paddy leaf detection using image processing" *Journal of global research in computer science ISSN:* 2229-371X, Volume-3, No 10, October 2012

[9]. Agriculture in India from Wikipedia, the free encyclopedia

[10]. Digital image processing from Wikipedia, the free encyclopedia.